

CORRESPONDENCE

Open Access

Human genetic research, race, ethnicity and the labeling of populations: recommendations based on an interdisciplinary workshop in Japan

Yasuko Takezawa^{1*}, Kazuto Kato², Hiroki Oota³, Timothy Caulfield⁴, Akihiro Fujimoto⁵, Shunwa Honda⁶, Naoyuki Kamatani⁷, Shoji Kawamura⁸, Kohei Kawashima⁹, Ryosuke Kimura¹⁰, Hiromi Matsumae³, Ayako Saito¹¹, Patrick E Savage¹², Noriko Seguchi¹³, Keiko Shimizu¹⁴, Satoshi Terao¹⁵, Yumi Yamaguchi-Kabata¹⁶, Akira Yasukouchi¹⁷, Minoru Yoneda¹⁸ and Katsushi Tokunaga¹⁹

Abstract

Background: A challenge in human genome research is how to describe the populations being studied. The use of improper and/or imprecise terms has the potential to both generate and reinforce prejudices and to diminish the clinical value of the research. The issue of population descriptors has not attracted enough academic attention outside North America and Europe. In January 2012, we held a two-day workshop, the first of its kind in Japan, to engage in interdisciplinary dialogue between scholars in the humanities, social sciences, medical sciences, and genetics to begin an ongoing discussion of the social and ethical issues associated with population descriptors.

Discussion: Through the interdisciplinary dialogue, we confirmed that the issue of race, ethnicity and genetic research has not been extensively discussed in certain Asian communities and other regions. We have found, for example, the continued use of the problematic term, "Mongoloid" or continental terms such as "European," "African," and "Asian," as population descriptors in genetic studies. We, therefore, introduce guidelines for reporting human genetic studies aimed at scientists and researchers in these regions.

Conclusion: We need to anticipate the various potential social and ethical problems entailed in population descriptors. Scientists have a social responsibility to convey their research findings outside of their communities as accurately as possible, and to consider how the public may perceive and respond to the descriptors that appear in research papers and media articles.

Keywords: Population descriptors, Ethics, Labeling, Race, Ethnicity, Japanese, Asian, Mongoloid

Background

With the rapid technical advances that have occurred in genome research, human genetic samples can now be analyzed on a massive scale and at an unprecedented speed. It is likely only a matter of time before this avalanche of genomic information is harnessed to allow healthcare decisions, such as the use of pharmaceuticals and the stratification of treatment protocols, to be increasingly tailored in a manner that will be informed by individual genetic predispositions. There are, of course,

many social and ethical issues involved in human genome research and in the application of the emerging knowledge [1,2]. One of the important, yet potentially complex, issues is how best to describe and report the populations that are being studied in the exploration of genetic variations [3]. There is concern that the use of improper and/or imprecise terminology, particularly language tied to concepts of "race" and "ethnic group" has the potential to both generate and reinforce racial and ethnic prejudices and diminish the clinical value of relevant research, as the massive literature shows [4,5]. In addition, broader terms such as "continental ancestry group" may not satisfactorily capture population differentiation on a sub-continental scale [6].

Full list of author information is available at the end of the article



^{*} Correspondence: yasuko@zinbun.kyoto-u.ac.jp

¹Institute for Research in Humanities, Kyoto University, Yoshidahonmachi, Sakyo-ku, Kyoto 606-8501, Japan

The issue of population descriptors has attracted a good deal of academic attention in North America and some European communities [7]. In these regions, there is much sensitivity to the ways populations are described. This is likely due, in part, to the history of racism within biomedical research and growing social awareness of the significance of ethnic and racial issues.

In contrast, many other regions, including Japan and some other regions in Asia, where the myth of raceless society have long persisted, have failed to tackle the topic, resulting in population descriptors sometimes being overgeneralized and ethically problematic. The number of genomic studies is skyrocketing in many countries and it is urgent that researchers based outside Europe and North America take a more active role in addressing the issues associated with the use and misuse of population descriptors.

Recognizing the importance of addressing these issues, we held a two-day workshop on January 7-8th, 2012, in Tokyo, which brought together scholars in diverse fields including those in the humanities, social sciences, medical sciences and genetics. The scientists shared their actual practices and relevant experiences on the use of population descriptors in publications and communications including review processes, while researchers in the humanities and social sciences discussed racism in the past and contemporary social issues involving minority groups. Although our focus was on the social and ethical issues involving Asian communities, in particular Japan, we hoped that our general conclusions would apply to other countries and communities with similar situations. At the end of the event, we agreed to produce a set of guidelines for reporting genetic studies involving populations in Asia, based in part on the recommendations of Caulfield et al. [8]. With this conference as the start, the discussion continued after the workshop through various methods such as core authors' meetings, e-mail communications, and video conferences. In this article, we report on the substance of these discussions focused on population descriptors and present the recommendations primarily targeting the genetics researchers in the region.

Discussion

Ethical and social issues arising from genetic research involving human populations

Research findings are often represented with over generalized descriptors such as "Asian" or other continental terms. In reality, samples are taken from much more discrete groups or specific and identifiable geographical regions. This tendency may be the result of a number of forces. Researchers sometimes attempt to draw more general conclusions than the actual data can support. In other cases, researchers may feel that without broad terms, it is difficult to gain recognition in the review process for publication or for the obtainment of a grant. The lack of education and

training for scientists and researchers regarding the use of descriptors and associated problems seems to be another cause of overgeneralization. Indeed, at the workshop, some of the participants shared their experiences of receiving such pressure to generalize from both research institutions and publishers. At the current time there is no data that maps the extent of this phenomenon and, as such both quantitative and qualitative systematic investigation is needed. The following examples demonstrate how, from the perspective of genetic research done in Asia, the inappropriate use of population descriptors could cause confusion and social controversies.

Although the term "Mongoloid" is rarely used today in North America and Europe, the situation is different in Japan and some other regions of Asia. Our preliminary analysis showed 113 hits in PubMed that contain the term "Mongoloid" in titles or abstracts of papers published during the period of 2004-2013, with no signs that use is decreasing. However, even among researchers, there is little awareness of the issues and little consistency in use, and its meaning can vary significantly depending on context [9-11]. Some researchers may use the term to designate a population in a particular or a variety of regions, including Eastern Asia, Southeast Asia or indigenous peoples in North America [12]. For others, it refers only to East Asians, or may be a synonym for the more generic Asian [13,14]. Moreover, the term has, in the past, been used to refer to individuals with Down's syndrome. In general, despite its continuing use, the term is problematic both because of the uncertainty regarding the population referred to, and because of its past controversial use [15,16].

Another example of the challenges associated with the use of population descriptors can be found in the frequent use of the terms European, African, and Asian. These continental terms are tremendously broad in scope. At the Tokyo meeting, for example, it was noted that even among the Japanese researchers, there was no unitary understanding of what populations should be considered "Asian."

More importantly, these terms can, in some contexts, be interpreted as referring to white, black, and Asian, the three classic, and socially constructed "races." There continues to be a great deal of academic work that highlights the degree to which these broad "racial" categories are, in reality, social constructs [17-19]. Although we should not overlook the correlation between "race" and socio-economic inequality involving factors such as health care and medical care, such discussion has usually arisen within the context of some North American and European societies. However, outside of these societies, the divergence between samples and population descriptors is also problematic. When the actual samples in the name of "European", "African", and "Asian" are taken from certain limited groups, without taking into account significant diversity within each region, it is unlikely that such broad terms have any scientific meaning, at least from the perspective of genetics on the global level [20,21]. Moreover, the research results may be taken as supporting the classic "racial" categories, with any discovered "differences" misinterpreted as genetically determined "racial differences."

The importance of the distinction between race and ethnicity cannot be overemphasized as the latter pays close attention to (presumably) shared cultural factors such as language, diet, and religion [22]. When considering the contribution of environmental as well as genetic factors to diversity within each continental region, the scientific validity of the use of such broad terms to describe samples becomes even more questionable.

In contrast to the above tendency to prefer broad terms, an influential study based on genome-wide 50 K SNP data reveals the detailed patterns of genetic differentiations within "Asians" [23]. The genetic ancestry of most populations was associated with ethnic and linguistic affiliations. Along the same lines, an analysis of 7,003 individuals from across Japan reveals interesting regional variations within the "Japanese" population. At one level, most Japanese fell into two main clusters from individuals taken in mainland Japan and those in Okinawa in a principal component analysis (PCA) plot based on genome-wide 140 K SNP data. Upon closer look, even among mainland Japanese, statistically meaningful genetic differentiation was found among individuals in different regions, such as Tohoku, Kanto, Kinki, and Kyushu [24].

The above study highlights that even populations traditionally presumed to have a high degree of homogeneity may have local genetic differentiations, that make the use of broader population terms less scientifically or clinically relevant. Researchers should strive to select terms that, as much as possible, reflect the sample population and nature of each study. Since genetic subpopulation structure is still generally unknown, sampling without considering the specifics of the subject population could cause false positive results on risk alleles of diseases. In addition, differences in whole genome sequences between individuals belonging to different populations should not be overgeneralized and misinterpreted as population differences.

Through our dialogue, it became apparent that the ways in which descriptors are selected sometimes differ depending on specialized fields. For example, researchers in physical/biological anthropological studies have a relatively long history of working on population genetics studies concerning local residents from whom they obtain sample data, and accumulate information on various populations from the perspective of long-term human evolution. Medical studies, on the other hand, are more concerned with the applicability of genetic studies contributing to the diagnosis and treatment of diseases. Disease gene surveys often take samples from patients at hospitals without controlling such factors as current location of residence or

generational continuity in each place. Such disciplinary differences in research purposes and methods have sometimes created different understandings and placed varying levels of attention on the issue of population description. This is one example why dialogue between scholars in different disciplines is indispensable in considering appropriate population descriptors.

There has been a growing discussion of the "co-production" of knowledge by the interplay between science and society [25]. The popular press is often blamed for the use of inappropriate or imprecise terms in the context of population genetic studies, whereas many scientists may believe that they take adequate precautions when describing the study samples, defining populations, and presenting discussions based on their research results. However, evidence indicates that imprecise and less than ideal descriptors are introduced throughout the research communication process [26]. If these descriptors are not carefully chosen, they create the potential for confusion both within the scientific community and in the wider society, leading to research inefficiencies and various social, ethical, and clinical problems [7].

What, then, would be a more desirable way to describe populations under study? The key is to use population descriptors that are scientifically valid for the particular study. For the first step, we recommend the use of population descriptors with more specific characteristics, such as geographical location and ethnic labeling as previously attempted - albeit imperfectly [27] - by research initiatives like the International HapMap Project [28]. This recommendation is based on the fact that various studies demonstrate the strong correlation between genetic distances and distances based on geography as well as ethnic affiliations [23,29]. This is, we believe, a better solution, but not a final one. Even when scientists choose more specific terminology, they have to explain the rationale behind the descriptors and what rules they employ in selecting the samples and defining the population.

Finally, the importance of education for undergraduate and graduate students as well as young trainees in human genetics and medicine cannot be overemphasized. It is urgent to prepare appropriate curriculums incorporating these ethical and social issues in order to effectively change the awareness of scholars and practitioners in the near future.

Recommendations

Based on the discussions and analyses described in the previous section, we have come up with the following nine recommendations.

1. In selecting descriptors, use specific names for populations or groups of people closely reflecting the make-up of the sample, while protecting the privacy of individuals included.

- 2. When using ethnological information for labeling, respect the cultural sensitivities of the populations and employ names that correspond to their cultural and ethnic backgrounds as much as possible. If not, clarify the definition of the names of populations used in the study, and explain why such descriptors have been chosen.
- 3. Explain how, where, and when sample data are collected, and who the concerned individuals are donors—as long as the information does not impinge on individual privacy. Also, the description of sampling date is important because allele frequencies could change in a population owing to demography, migration, and drift in a short time span.
- 4. Avoid overly broad category names such as Asian, European, or African. Recognize that the use of such names without scientific justification could cause confusion, misinterpretation, and social controversy particularly if the research results are interpreted in a manner that could emphasize the existence of these "racial" categories. If the use of broad category names is necessary, there must be sufficient scientific grounds and explanation.
- 5. When genetic differences are ones in degrees of frequencies among different populations, avoid typological discussions and emphasize that the differences are a matter of frequency and probability, not differences that are clear-cut and discrete.
- 6. Be alert to the possibility that research on human populations could cause various kinds of social and ethical problems; therefore, endeavor to take steps to anticipate relevant issues that may emerge, including seeking to collaborate, as appropriate, with colleagues in other relevant disciplines (e.g., medical researchers and researchers in the humanities and social sciences). It is also important to recognize that scientific activities are also influenced by social and political factors. Population descriptors are no exception.
- 7. Prepare an easily understandable summary of research findings, so that reporters for newspapers and other popular media can prepare proper reports that utilize appropriate population descriptors.
- 8. Point out any mistakes or misinterpretations of the research results after the reports are released to the public, and if opportunity allows, confirm them before public release.
- 9. Incorporate the above considerations into the education curriculums for emerging trainees at an early stage in their careers [8].

Summary

In this age of genomics, differences between populations are often reported as having genetic bases [26]. However, misunderstanding and extended interpretation of the results

might contribute to discrimination, or justify health care and socio-economic inequalities [30]. Therefore, we need to anticipate the various potential social and ethical problems associated with population descriptors. Scientists have a social responsibility to convey their research findings outside of their communities as accurately as possible, and to consider how the public may perceive and respond to the descriptors that appear in research papers and media articles. Researchers in the humanities and social sciences may be able to contribute to the identification of potential social and ethical problems involving population descriptors. As such, there is a compelling need for truly interdisciplinary dialogue and collaboration between professionals in genetics, medical science, the humanities, and social sciences. We believe that such activities are particularly needed in the countries and communities outside of North America and Europe, where the issues of race, ethnicity and genetic research are not discussed extensively.

What we have discussed is merely a first step in the process of addressing the challenges associated with the use of population descriptors in the context of genetic research, and we hope that it will encourage action and the exchange of ideas and opinions, especially among the relevant research communities in Japan and other Asian countries.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YT, KK², HO conceived the idea of the research and organized the workshop in Tokyo in 2012. YT, KK, HO, TC and KT wrote the core of the manuscript. TC, NK, AF, NS, YYK, KT made presentations at the conference. HM conducted the preliminary analyses for the discussion. SH, SK, KK⁹, RK, AS, PES, KS, ST, AY, MY contributed to discussion at the workshop and subsequently gave comments on the drafts of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Prior to the two-day conference, "The Interface of the Humanities and Genomics, Part II", Part I was held in January 2011 in Kyoto, and another conference focusing on "Okinawans" was held in March 2011 in Okinawa. These events, as well as the one in Tokyo in 2012, are part of the collaborative research project, "A Japan-Based Global Study of Racial Representations," the Grant-in-Aid for Scientific Research (S) No. 22222003 (principal investigator: Yasuko Takezawa), generously funded by the Japan Society for the Promotion of Science. We would like to extend our appreciation to all other scholars including those who came from abroad to Japan to take part in previous events and contribute to these discussions. We also thank Chiori Goto and Yuka Kanno for their editorial support as well as participation at the conference and Mika Ko and Wataru Kusaka for participation at the workshop.

Author details

¹Institute for Research in Humanities, Kyoto University, Yoshidahonmachi, Sakyo-ku, Kyoto 606-8501, Japan. ²Department of Biomedical Ethics and Public Policy, Graduate School of Medicine, Osaka University, Suita, Japan. ³Laboratory of Genome Anthropology, Department of Anatomy, Kitasato University School of Medicine, Sagamihara, Japan. ⁴Faculty of Law and the School of Public Health, University of Alberta, Edmonton, Canada. ⁵Laboratory of Medical Informatics, RIKEN Center for Genomic Medicine, Yokohama, Japan. ⁶The Open University of Japan, Tokyo, Japan. ⁷Institute of Data Analysis, StaGen Co.Ltd., Tokyo, Japan. ⁸Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo,

Kashiwa, Japan. ⁹Department of English Language and British and American Studies, Faculty of Humanities, Musashi University, Tokyo, Japan. ¹⁰Department of Human Biology and Anatomy, Graduate School of Medicine, University of the Ryukyus, Nishihara, Japan. ¹¹Department of Art Studies, Meiji Gakuin University, Tokyo, Japan. ¹²Department of Musicology, Tokyo University of the Arts, Tokyo, Japan. ¹³Department of Environmental Changes, Graduate School of Social and Cultural Studies, Kyushu University, Fukuoka, Japan. ¹⁴Department of Zoology, Faculty of Science, Okayama University, Kobe, Japan. ¹⁶Department of Integrative Genomics, Tohoku Medical Megabank Organization, Tohoku University, Miyagi, Japan. ¹⁷Department of Human Science, Faculty of Design, Kyushu University, Fukuoka, Japan. ¹⁸Laboratory for Radiocarbon Dating, The University Museum, The University of Tokyo, Tokyo, Japan. ¹⁹Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan.

Received: 18 March 2013 Accepted: 11 April 2014 Published: 23 April 2014

References

- Clayton EW, Smith M, Fullerton SM, Burke W, McCarty CA, Koenig BA, McGuire AL, Beskow LM, Dressler L, Lemke AA, Ramos EM, Rodriguez LL: Confronting real time ethical, legal, and social issues in the Electronic Medical Records and Genomics (eMERGE) consortium. Genet Med 2010, 12:616–620.
- Meslin EM, Cho MK: Research ethics in the era of personalized medicine: updating science's contract with society. Public Health Genomics 2010, 13:378–384.
- 3. Ali-Khan S, Krakowski T, Tahir R, Daar A: The use of race, ethnicity and ancestry in human genetic research. *HUGO J* 2011, 5:47–63.
- Dawson G: Human genome, race and medicine. J Natl Med Assoc 2003, 95:309–312.
- Lee SSJ, Mountain J, Koenig B, Altman R, Brown M, Camarillo A, Cavalli-Sforza L, Cho M, Eberhardt J, Feldman M, Ford R, Greely H, King R, Markus H, Satz D, Snipp M, Steele C, Underhill P: The ethics of characterizing difference: guiding principles on using racial categories in human genetics. Genome Biol 2008, 9:1–4.
- Aspinall PJ: The operationalization of race and ethnicity concepts in medical classification systems:issues of validity and utility. Health Informat J 2005. 11(4):259–274.
- Rugnetta M, Desai K: Addressing race and genetics health disparities in the age of personalized medicine. http://scienceprogress.org/2011/06/addressingrace-and-genetics/ (accessed 13 March 2014).
- Caulfield T, Fullerton S, Ali-Khan S, Arbour L, Burchard E, Cooper R, Hardy B-J, Harry S, Hyde-Lay R, Kahn J, Kittles R, Koenig B, Lee S, Malinowski M, Ravitsky V, Sankar P, Scherer S, Séguin B, Shickle D, Suarez-Kurtz G, Daar A: Race and ancestry in biomedical research: exploring the challenges. Genome Med 2009, 1:1–8.
- Terada T, Kaneko H, Li AL, Kasahara K, Ibe M, Yokota S, Kondo N: Analysis of Ig subclass deficiency: first reported case of IgG2, IgG4, and IgA deficiency caused by deletion of C alpha 1, psi C gamma, C gamma 2, C gamma 4, and C epsilon in a Mongoloid patient. J Allergy Clin Immunol 2001, 108:602–606.
- Park TS, Oh SH, Choi JC, Lee DD, Kim HH, Chang CL, Lee EY, Son HC: The clinical significance of antibody screening test including Dia + panel cell in Asian-Mongoloid populations. J Korean Med Sci 2003, 18:669–672.
- Wey MC, Shim CN, Lee MY, Jamaluddin M, Ngeow WC: The safety zone for mini-implant maxillary anchorage in Mongoloids. Aust Orthod J 2012, 28:17–21.
- Tokunaga K, Imanishi T, Takahashi K, Juji T: On the origin and dispersal of East Asian populations as viewed from HLA haplotypes. In *Prehistoric Mongoloid dispersals*. 1st edition. Edited by Akazawa T, Szathmáry EJE. New York: Oxford University Press; 1996:187–197.
- Oota H, Kitano T, Jin F, Yuasa I, Wang L, Ueda S, Saitou N, Stoneking M: Extreme mtDNA homogeneity in continental Asian populations. Am J Phys Anthropol 2002, 118:146–153.
- Oota H, Saitou N, Ueda S: A large-scale analysis of human mitochondrial DNA sequences with special reference to the population history of East Eurasian. Anthropol Sci 2002, 110:293–312.
- Brace CL: "Race" is a four-letter word: the genesis of the concept. New York: Oxford University Press; 2005.

- 16. Takezawa Y: Problems with the terms: "Caucasoid", "Mongoloid" and "Negroid". Zinbun 2011, 43:61–68.
- Smart A, Tuton R, Martin P, Ellison GTH: "Race" as a social construction in genetics. In Identity politics and the new genetics: re/creating categories of difference and belonging. Edited by Schramm K, Skinner D, Rottenburg R. New York: Berghahn Books; 2012:30–52.
- 18. Hartigan J: Is race still socially constructed? The recent controversy over race and medical genetics. *Sci Cult* 2008, 17(2):163–193.
- Reardon J: Decoding race and human difference in a genomic age. Differ J Feminist Critic Stud 2004, 15(3):38–65.
- Duster T: The role of molecular genetics in the shifting boundaries of human taxonomies. In Racial representations in Asia. Edited by Takezawa Y. Kyoto: Kyoto University Press; 2011:188–204.
- Duster T: Social diversity in humans: implications and hidden consequences for biological research. In Human Variation: A Genetic Perspective on Diversity, Race, and Medicine. Edited by Chakravarti A. Cold Spring Harbor, NY: Cold Spring Harbor Press; 2014.
- Lee C: "Race" and "ethnicity" in biomedical research: how do scientists construct and explain differences in health? Soc Sci Med 2009, 68:1183–1190.
- HUGO Pan-Asian SNP Consortium, Abdulla MA, Ahmed I, Assawamakin A, Bhak J, Brahmachari SK, Calacal GC, Chaurasia A, Chen CH, Chen J, Chen YT, Chu J, la Paz EM C-d, De Ungria MC, Delfin FC, Edo J, Fuchareon S, Ghang H, Gojobori T, Han J, Ho SF, Hoh BP, Huang W, Inoko H, Jha P, Jinam TA, Jin L, Jung J, Kangwanpong D, Kampuansai J, et al: Mapping human genetic diversity in Asia. Sci 2009, 326:1541–1545.
- Yamaguchi-Kabata Y, Nakazono K, Takahashi A, Saito S, Hosono N, Kubo M, Nakamura Y, Kamatani N: Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. Am J Hum Genet 2008, 83:445–456.
- 25. Jasanoff S: States of knowledge: the co-production of science and social order. London, New York: Routledge; 2004.
- Rachul C, Ouellette C, Caulfield T: Tracing the use and source of racial terminology in representations of genetic research. *Genet Med* 2011, 13:314–319.
- Institute of Medicine, Smedley BD, Sith AY, Nelson AR: Unequal Treatment: Confronting Racial and Ethnic Disparities in Health. Washington, D.C: Institute of Medicine; 2002.
- 28. The International HapMap C: A haplotype map of the human genome. *Nature* 2005, **437**:1299–1320.
- Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, Bustamante CD: Genes mirror geography within Europe. Nature 2008, 456:98–101.
- Hamilton JA: Revitalizing difference in the HapMap: race and contemporary human genetic variation research. J Law Med Ethics 2008, 36(3):471–477.

doi:10.1186/1472-6939-15-33

Cite this article as: Takezawa *et al.*: Human genetic research, race, ethnicity and the labeling of populations: recommendations based on an interdisciplinary workshop in Japan. *BMC Medical Ethics* 2014 **15**:33.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit

